Am J Hum Genet 35:1139–1155, 1983

The Search for Heterogeneity in Insulin-Dependent Diabetes Mellitus (IDDM): Linkage Studies, Two-Locus Models, and Genetic Heterogeneity

SUSAN E. HODGE,[1] CAROL E. ANDERSON,[2] KATHERINE NEISWANGER,[3] ROBERT S. SPARKES,[3,4] AND DAVID L. RIMOIN[5]

SUMMARY

One hundred families with insulin-dependent diabetes mellitus (IDDM) were analyzed for linkage with 27 genetic markers, including HLA, properdin factor B (BF), and glyoxalase 1(GLO) on chromosome 6, and Kidd blood group (Jk) on chromosome 2. The linkage analyses were performed under several different genetic models. An approximate correction for two-locus linkage analysis was developed and applied to four markers. Two different heterogeneity tests were implemented and applied to all the markers. One, the Predivided-Sample Test, utilizes various criteria thought to be relevant to genetic heterogeneity in IDDM. The other, the Admixture Test, looks for heterogeneity without specifying a priori how the sample should be divided.

Results continued to support linkage of IDDM with three chromosome 6 markers: HLA, BF, and GLO. The total lod score for Kidd blood group, under the recessive model with 20% penetrance, is 1.63—down 1.2 from the 2.83 reported by us earlier. The only other marker whose lod score exceeded 1.0 under any model was pancreatic amylase (AMY2). The two-locus correction, which involved lowering the penetrance values used in the analysis, affected estimates of θ (recombination fraction) but did not markedly change the lod scores themselves. There was little evidence for heterogeneity within any of the lod scores, under either the Predivided-Sample Test or the Admixture Test.

Received December 27, 1982; revised March 2, 1983.

This work was supported by grants AM-25834 and MH-33617 from the National Institutes of Health, a grant from the Kroc foundation, and NIH grant HD-04612 from the UCLA Mental Retardation Research Center. Computer assistance was obtained from the UCLA Office of Academic Computing and the Human Genetic Laboratory's PDP 11/44.

[1] Department of Biomathematics, UCLA School of Medicine, Los Angeles, CA 90024. Request reprints from: S. E. H., MRRC/NPI, 760 Westwood Plaza, UCLA School of Medicine, Los Angeles, CA 90024.

[2] Department of Pediatrics, University of California, Irvine, Calif.

[3] Department of Psychiatry, UCLA School of Medicine.

[4] Department of Medicine, UCLA School of Medicine.

[5] Departments of Pediatrics and Medicine, Harbor-UCLA Medical Center, Torrance, CA 90509.

INTRODUCTION

The mode of inheritance of insulin-dependent diabetes mellitus (IDDM) (type 1 diabetes) is not fully understood. Family and twin studies indicate that IDDM has a genetic component but that penetrance may be as low as 20% [1]. Several researchers have proposed genetic heterogeneity within IDDM [2–8]. Rotter [9] summarizes many of the heterogeneity hypotheses and the evidence supporting each.

Consistent associations observed between IDDM and alleles of the HLA and BF (properdin factor B) marker systems [10–13] may provide clues to the genetics of IDDM. These associations may imply that an IDDM susceptibility locus (or loci) exists on chromosome 6, tightly linked to HLA [14]. However, formal genetic linkage studies between IDDM and HLA should be interpreted with caution, since the association may confound the linkage results [15, 16].

Until recently, little attention had been paid to other genetic markers and IDDM. Two studies [17, 18] examined associations between IDDM and a variety of markers, with inconclusive results. We also examined associations in the current data set [19]. In a preliminary study [20] of 71 families (33 multiplex), we reported suggestive evidence for genetic linkage between IDDM and the Kidd (Jk) blood group on chromosome 2, as well as between IDDM and three chromosome 6 markers—HLA, BF, and GLO. We raised the possibility of two-locus inheritance for IDDM, with one locus linked to HLA and the other linked to Jk. Two-locus inheritance had already been suggested by other authors on mathematical grounds [21, 22], without any speculation as to where a second locus might be located.

Our current study uses linkage analysis to look for heterogeneity within IDDM. This paper is organized into three sections: part (1), *Linkage Update*, extends the lod score linkage analysis reported earlier [20] to include our complete data set of 100 IDDM families; part (2), *"Two-locus" Linkage Analysis/Reduced Penetrance*, develops an approximate correction for a two-locus linkage analysis by reducing the penetrance used in the linkage analysis to allow for effects of a possible second locus, and applies this correction to four of the markers; and part (3), *Heterogeneity*, implements two different methods of testing for genetic heterogeneity within lod scores and applies them to all the markers.

METHODS

Families were ascertained through at least one member affected with IDDM by age 40. Multiplex families are the most informative for linkage, so they were sought preferentially. One hundred families (50 multiplex, 11 multigenerational, and 39 simplex), containing a total of 560 individuals, were analyzed for linkage.* Most of the families were nuclear. Clinical details are given by Anderson et al. [23]. Available family members were typed for 27 red blood cell antigens, red blood cell enzymes, and serum protein markers,† using standard laboratory techniques. Not every individual was typed for every marker. Some

* The families were DI numbers 002–003, 005–016, 018–027, 029–050, 052–058, 061–062, 065–098, 100–109, and 999 (UCLA linkage file).

† ABO, ACP1, ADA, AK1, AMY2, BF, CHE1, CHE2, Fy, ESD, GALT, GC, GLO, GPT, HLA, HPA, Jk, K, Le, MNSs, P, PI, PGD, PGM1 (subtyped), PGP, Rh, and TF. Gene symbols are those used by the Human Gene Mapping Conference [24].

individuals were typed for Jk^a only, because of a temporary lack of Jk anti-b antisera. We described the distributions of the markers among unrelated IDDM patients, one from each family, in detail [19].

(1) *Linkage Update*

Linkage was analyzed by the method of lod scores [25], as calculated by the computer program LIPED [26] and modified to allow for variable age of onset [27]. We examined lod scores at a 7 × 7 matrix of recombination fractions for the two sexes: θ(male), θ(female) = .01, .05, .10, .20, .30, .40, and .50. A computer program (QUAD) was written to interpolate the approximate maximum lod score and maximum likelihood estimator (MLE) of θ, using quadratic interpolation as suggested by Rao et al. [28]. The quadratic interpolation was generalized to the two-dimensional case by fitting the best least-squares paraboloid to the nine points around and including the calculated maximum lod score. HLA was coded by haplotype [29], with recombinant individuals omitted from the analysis.

To generate lod scores, one must specify the mode of inheritance of the disease at the locus being investigated for linkage. Since the genetics of IDDM are unclear, we used three genetic models—recessive with 20% penetrance (model 1), recessive with 50% penetrance (model 2), and dominant with 20% penetrance (model 3). Let the presumed IDDM locus have two alleles D and d, with gene frequencies p and q. The diabetogenic allele is d under models 1 and 2, and D, under model 3. Assuming an overall population prevalence $K = .005$ [30], we used the following gene frequencies: $q = .158$ for model 1, $q = .10$ for model 2, and $p = .0126$ for model 3. A linear age-of-onset penetrance curve goes from 0% penetrance at birth to either 20% (models 1 and 3) or 50% penetrance (model 2) at age 40.

Four markers, HLA, BF, GLO, and Jk, yielded interesting lod scores under these first three models. Model 4, based on an intermediate mode of inheritance proposed by Spielman et al. [30], was then used to analyze these four markers further. This model allows for differential penetrances, as a function of gene dosage. Let d be the diabetogenic allele. Then the dd genotype has a maximum penetrance of .662 at age 40, and the Dd genotype, .062; the frequency of allele d is $q = .035$.

(2) *"Two-locus" Linkage Analysis/Reduced Penetrance*

Assume that there are two disease loci, for example, one linked to HLA and the other linked to Jk, such that the IDDM phenotype depends on the genotypes at both loci. A linkage analysis with the chromosome 6 markers should allow for the effects of the "Jk-linked" disease locus, and vice versa for a linkage analysis with Jk.

The major problem, for a clinically unaffected person, is: What is the genotype? Linkage analysis uses penetrance to answer this question. Our correction is based on two probabilities:

TABLE 1

POPULATION PREVALENCE K AND PENETRANCES f_1 AND f_2 FOR THREE SIMPLE TWO-LOCUS MODELS

Model	K	f_1	f_2
R-R	q^2s^2F	s^2F	q^2F
D-R	$(1 - q^2)s^2F$	s^2F	$(1 - q^2)F$
D-D	$(1 - q^2)(1 - s^2)F$	$(1 - s^2)F$	$(1 - q^2)F$

NOTE: See text for description of the three models. q = gene frequency at locus 1; s = gene frequency at locus 2. F = overall penetrance. D-R model is shown for locus 1 dominant, locus 2 recessive.

TABLE 2

SINGLE-LOCUS ANALYSES USED IN "TWO-LOCUS" LINKAGE ANALYSIS

MODEL	CASE	LOCUS 1		LOCUS 2	
		f_1	q	f_2	s
		Recessive		Recessive	
R-R	1.....	.20*	.158*	.005†	1.0†
	2.....	.032	.398	.032	.398
		Dominant		Recessive	
D-R	1.....	.20*	.9874*	.005†	1.0†
	2.....	.032	.9175	.032	.398
	3.....	.005†	0†	.20*	.158*
		Dominant		Dominant	
D-D	1.....	.20*	.9874*	.005†	0†
	2.....	.032	.9175	.032	.9175

NOTE: D-R model is shown for locus 1 dominant, locus 2 recessive; q and s are gene frequencies, f_i = penetrance at locus i.

* These values correspond to a single-locus analysis performed earlier.

† These values correspond to a ubiquitous locus and yield no linkage information.

(1) the probability of an individual being affected, given that he has an at-risk genotype at *one* locus, and (2) the probability of being affected, given that his genotype at *both* loci puts him at risk. In a two-locus disease, the first probability is much lower than the second. The second probability is what we come closest to observing directly, for example, via the monozygotic (MZ) twin concordance rate. Yet, the first probability is what we need for a linkage analysis. Therefore, we allowed for second-locus effects by reducing the penetrance at the locus being analyzed, then performing the analysis as if for a single-locus trait. The APPENDIX gives mathematical details and an example.

Table 1 describes the parameters of three basic two-locus models: recessive-recessive (R-R), dominant-recessive (D-R), and dominant-dominant (D-D), as developed in [31] but modified here to allow for reduced penetrance. For each model, selected sets of parameter values were chosen for analysis, as described in the APPENDIX. Table 2 shows the single-locus analyses we performed, corresponding to these selected cases.

(3) *Heterogeneity*

Predivided-Sample Test. We searched for linkage heterogeneity in two ways. In the first, the Predivided-Sample Test, four analyses were performed, each using a different criterion to divide the families into two groups. (Families not falling unambiguously into either of the two groups were not included in that particular analysis.) The four analyses were: (A) *Simplex vs. multiplex families.* Here, group 1 consisted of simplex families, defined as having only one affected child, plus at least one unaffected sibling [23]. Group 2 consisted of multiplex families with at least two affected children within a nuclear family, with or without an affected parent. (B) *HLA-associated haplotypes.* Group 1, the "associated-haplotype" group, consisted of nuclear families in which anyone—proband, parents, or siblings—had any of the haplotypes *B8-DR3*, *B18-DR3*, *B15-DR4*, or *B62-DR4*.* If no one in the nuclear family had any of these haplotypes, the family was put in

* B62 is a split of B15 [32].

the "not-associated" group—group 2. (C) HLA-DR3/4 *index cases*. Families were divided according to whether the oldest affected child had the *DR3/4* genotype (group 1) or not (group 2). (D) *Positive vs. negative Jk lod scores*. Jk lod scores under the recessive, 20% model (model 1) were examined at θ(male) = .10, θ(female) = .10. Families were assigned to group 1 or 2 depending on whether these lod scores were $\geq .1$ or $\leq -.1$.

In all four analyses (A)–(D), we tested for heterogeneity by looking for significant differences between groups 1 and 2. Let $\hat{z}_i$ = maximum lod score for group i (i = 1.2) and $\hat{Z}$ = maximum lod score for the two groups merged together. Then $\chi^2 = 2(\ln 10)(\hat{z}_1 + \hat{z}_2 - \hat{Z})$ is asymptotically distributed as a chi square with 1 df [33, 34]. This χ^2 provides a heterogeneity test statistic.

Admixture Test. The second approach to linkage heterogeneity, the Admixture Test, followed a suggestion by Smith [35]. This method has also been used by Ott [36] and applied by Risch and Baron [37] to bipolar affective disorder. Instead of dividing the families into two groups according to a predetermined criterion, assume that some unspecified proportion α of families are linked to the marker with recombination fraction θ, whereas the remaining $(1 - \alpha)$ are unlinked. The likelihood of the sample, $L(\theta,\alpha)$, can be expressed as $L(\theta,\alpha) = \alpha L(\theta) + (1 - \alpha)L(0.5)$. [Let "$L(\theta)$" denote $L(\theta,1)$.] On the right side of the equation above, $L(\theta)$ represents the likelihood used in conventional lod scores, with α = 1; and α functions as a prior probability for each family to be in the "linked," as opposed to the "unlinked," group. $L(\theta,\alpha)$ can be maximized with respect to θ and α simultaneously. The Admixture Test is applied in one of two ways, depending on whether or not we have prior evidence that the disease and marker are linked.

(A) *Prior knowledge of linkage—heterogeneity statistic "H."* In this situation, assume linkage exists, and test only for heterogeneity. The null hypothesis is H_0: $\alpha = 1$. The likelihood ratio (LR) is:

$$LR = \frac{\max_{\theta,\alpha} L(\theta,\alpha)}{\max_{\theta} L(\theta)} .$$

Define the heterogeneity statistic $H = 2(\ln LR)$; then H is asymptotically distributed as chi square with 1 df [34].

(B) *No prior knowledge of linkage—heterogeneity statistic "lod2."* To test simultaneously for heterogeneity *and* linkage, use H_0: $\theta = .5$, $\alpha = 1$. The appropriate test statistic is the log LR, or modified lod score, denoted "lod2":

$$\text{lod2} = \log_{10} \frac{\max_{\theta,\alpha} L(\theta,\alpha)}{L(0.5)} .$$

As with standard lod scores, a value of 3.0 is required for statistical significance. Note that the standard lod score, Z, is related to lod2 and H as follows:

$$Z = \log_{10} \frac{\max_{\theta} L(\theta,1)}{L(0.5)} = \text{lod2} - \frac{H}{2(\ln 10)} .$$

Also note that the maximum likelihood estimate (MLE) of θ from the "lod2" score will not, in general, equal $\hat{\theta}$ from the standard lod score.

We used the H score to test for linkage heterogeneity in the four markers already thought to be linked to IDDM (HLA, BF, GLO, and Jk) and the "lod2" score for all other markers. All scores were calculated at $\theta = \theta$(male) = θ(female) = .01, .05, .10, .20, .30, .40, and at α = 0, .05, .10, . . . , 1.0.

TABLE 3

SUMMARY OF LINKAGE ANALYSIS FOR RECESSIVE AND DOMINANT MODELS WITH 20% PENETRANCE

CHROMOSOME	MARKER SYSTEM		RECESSIVE (MODEL 1) Lod	RECESSIVE (MODEL 1) θ	DOMINANT (MODEL 3) Lod	DOMINANT (MODEL 3) θ	NO. INFORMATIVE FAMILIES
1	PGD		0.71	.01,.01	...*	...	16
	Rh		−7.56	.01,.01	−4.53	.01,.01	82
	PGM1	min:	−3.89	.01,.01	−1.61	.01,.40	71
		max:	0.55	.30,.20	1.02	.30,.05	
	AMY2.....		1.24	.01,.05	0.87	.01,.01	21
	Fy	min:	−4.70	.01,.01	−1.76	.01,.01	84
		max:	...	...	0.52	.10,.40	
2	Jk.....	min:	...	...	−1.48	.05,.40	78
		max:	1.63	.10,.10	1.23	.40,.01	
	ACP1		−4.90	.01,.01	−3.81	.01,.10	78
4	GC.....		−5.50	.01,.01	−2.05	.01,.01	73
	MNSs		−10.11	.01,.01	−5.54	.01,.01	80
6	HLA		4.77	.10,.10	6.17	.10,.01	100
	BF.....		2.90	.01,.10	3.74	.01,.01	61
	GLO	min:	−1.41	.01,.01	−1.62	.01,.01	94
		max:	1.16	.10,.20	0.88	.40,.05	
9	GALT		−0.85	.01,.01	...*	...	40
	ABO	min:	−3.25	.01,.01	−1.57	.40,.01	73
		max:	0.55	.20,.40	...	...	
	AK1		−1.27	.01,.01	...*	...	24
13	ESD.....		−2.23	.01,.01	−3.04	.01,.01	38
16	HPA.....		−9.32	.01,.01	−4.56	.01,.01	76
	PGP.....		−2.48	.01,.01	−2.35	.01,.01	35
Unassigned	K		−2.21	.01,.01	−1.56	.01,.01	32
	P	min:	−1.87	.01,.01	−0.52	.01,.40	81
		max:	0.84	.40,.01	0.53	.40,.01	
	GPT.....		−6.88	.01,.01	−4.24	.01,.01	80

NOTE: Minimum lod scores are shown only if ≤ -0.5; maximum lod scores are shown only if ≥ 0.5. First value of θ is male; second is female.

* Under this model, all lod scores fell between −0.5 and 0.5.

RESULTS

(1) *Linkage Update*

Table 3 summarizes the results of linkage analyses performed between IDDM and 21 informative markers, analyzed under recessive and dominant models with 20% penetrance (models 1 and 3).* Lod scores under the recessive model with 50% penetrance (model 2) were also run for all 27 markers. Results were qualitatively similar to those under model 1 and are not included in the table. Six markers—ADA, CHE1, CHE2, Le, PI, and TF—were uninformative, yielding both maximum and minimum lod scores between −0.5 and 0.5 under all three models. They are not included in the table.

Four markers—HLA, BF, GLO, and Jk—yielded maximum lod scores exceeding 1.0 under most of the three models (1–3). Of these, the HLA and BF scores are at or near statistical significance (lod = 3.0) under all three models. The Jk score

* Detailed lod scores for all 27 markers, under models 1, 2, and 3, are available upon request.

is over 1.6 for the two recessive models (1 and 2). A fifth marker, AMY2, has a maximum lod score of 1.24 under model 1 and 1.28 under model 2; under model 3, the maximum lod score was 0.87. Under some circumstances, close linkage could be ruled out for these markers as well: the HLA lod score at θ(male) = θ(female) = .01 was −4.53 under model 2; GLO lod scores at θ(male) = θ(female) = .01 ere −1.41 under model 1, −5.23 under model 2, and −1.62 under model 3; and the Jk lod score at θ(male) = .01, θ(female) = .40 was −1.48 under model 3.

The following markers yielded lod scores ≤ -1.0 under all or most of the three models and can probably be ruled out for close linkage with IDDM under any simple single-gene model: ACP1, Fy, ESD, GC, GPT, HPA, K, MNSs, P, PGM1, PGP, and Rh.

The four markers HLA, BF, GLO, and Jk were felt to be of greatest interest and were analyzed under model 4 (intermediate) as well. Table 4 summarizes results for these four markers under all four models, along the diagonal where θ(male) = θ(female). (The "two-locus model" columns in table 4 will be discussed below.) Both the maximum lod scores and the θ values under model 4 fall between or near those under models 1–3.

(2) *"Two-locus" Linkage Analysis/Reduced Penetrance*

Table 2 defined 14 separate single-locus analyses. However, many of these 14 analyses were duplicates of each other, were already performed above, or represented a ubiquitous locus (see the APPENDIX) and did not need to be done at all. Only two remained to be performed: recessive with penetrance $f_i = .032$, and dominant with $f_i = .032$. The "two-locus model" columns in table 4 give

TABLE 4

MAXIMUM LOD SCORE AND MLE OF θ (INTERPOLATED) FOR FOUR LINKAGE MARKERS AND IDDM. θ(MALE) = θ(FEMALE)

	RECESSIVE MODELS					
	"TWO-LOCUS" MODEL (f = .032)		MODEL 1 (f = .20)		MODEL 2 (f = .50)	
	$\hat{Z}$	$\hat{\theta}$	$\hat{Z}$	$\hat{\theta}$	$\hat{Z}$	$\hat{\theta}$
HLA.....	4.37	0(+.08)	4.89	.13 ± .06	4.40	.17 ± .07
BF.....	2.68	0(+0)	2.83	.07 ± .07	3.03	.11 ± .08
GLO.....	1.05	.06 ± .14	1.10	.19 ± .12	1.13	.22 ± .09
Jk	1.63	0(+0.8)	1.63	.10 ± .11	1.76	.13 ± .08

	DOMINANT MODELS				INTERMEDIATE MODEL	
	"TWO-LOCUS MODEL" (f = .032)		MODEL 3 (f = .20)		MODEL 4	
	$\hat{Z}$	$\hat{\theta}$	$\hat{Z}$	$\hat{\theta}$	$\hat{Z}$	$\hat{\theta}$
HLA.....	5.50	0(+0)	6.11	.06 ± .04	5.88	.07 ± .05
BF.....	3.08	0(+0)	3.81	0(+0)	3.63	0(+.07)
GLO.....	0.65	.12 ± .17	0.57	.20 ± .15	0.80	.18 ± .15
Jk	0.78	.05 ± .18	0.50	.17 ± .20	1.22	.08 ± .15

the results for the four markers HLA, BF, GLO, and Jk. As in the earlier analyses, the HLA and BF lod scores are statistically significant under both models, except the recessive with $f_i = .032$, where the maximum lod score for BF is 2.68 (interpolated). The lod scores with GLO and Jk are positive but not striking. Maximum likelihood estimates of θ are zero for HLA and BF; between 0 and .05 for Jk; and between .06 and .12 for GLO—lower than under the other models.

Thus, table 4 contains results for all single-locus analyses described in table 2. For example, consider the dominant-recessive (D-R) model, case 2, where the hypothesized chromosome 6 IDDM gene is recessive and the hypothesized Jk-linked gene is dominant, both with penetrance $f_i = .032$. Then the maximum lod scores between IDDM and the three loci HLA, BF, and GLO are 4.37, 2.68, and 1.05, respectively, with corresponding recombination fractions as shown in table 4; the maximum Jk lod score is 0.78, with $\theta = .05 \pm .18$.

(3) *Heterogeneity*

Predivided-Sample Test. Table 5 gives the asymptotic χ^2 values for the four Predivided-Sample analyses (A)–(D) performed on the four markers HLA, BF, GLO, and Jk. With one exception, none reaches the value of 3.84 required for statistical significance.

The very low χ^2 values for analysis (A), simplex vs. multiplex families, result from the fact that the simplex families yielded almost no linkage information. For all four markers, $\hat{z}_1 = 0$ with $\hat{\theta}_1 = .5$, and $\hat{z}_2 = \hat{Z}$.

When families are split by the presence or absence of a DR3/4 index case [analysis (C)], the GLO χ^2 (1 df) of 5.85 is statistically significant ($P < .05$), using the recessive 20% lod scores.

Admixture Test. In the Admixture Test, we calculated the H heterogeneity scores for HLA, BF, GLO1, and Jk, examined under models 1, 3, and 4 (recessive, 20% penetrance; dominant, 20%; and intermediate). In most cases, the maximum likelihood occurred at $\alpha = 1$ (i.e., no heterogeneity), and in no case was there statistically significant evidence for linkage heterogeneity under any of the three genetic models examined. The highest H value was 0.12, nowhere near the cutoff value of 3.84 required for significance at the 5% level.

TABLE 5

ASYMPTOTIC χ^2 VALUES (1 df) FOR FOUR HETEROGENEITY ANALYSES, USING THE PREDIVIDED-SAMPLE TEST

	MARKER			
ANALYSIS	HLA	BF	GLO	Jk
(A) Simplex vs. multiplex	0.11	0.31	0.02	0.23
(B) Associated haplotypes	2.73	1.29	2.88	0.74
(C) DR3/4 index case—model 1	0.11	0.35	5.85	0.04
—model 4	0.02	0.00	3.36	0.32
(D) Jk lod scores	1.61	0.90	0.67	...

NOTE: All lod scores are based on the recessive model with 20% penetrance (model 1) unless otherwise indicated.

"Lod2" scores were calculated for all other markers, under a recessive and a dominant model (models 1 and 3). For most markers, the maximum lod2 score occurred at $\alpha = 1.0$ and was therefore identical to the maximum standard lod score. For eight marker systems under the recessive model (seven under the dominant model), the maximum lod2 score occurred at $\alpha < 1$, but in no case did this score exceed 1.1. Thus, we found no statistically significant evidence for heterogeneity *and* linkage, using these markers.

DISCUSSION

(1) *Linkage Update*

The lod scores with HLA, BF, GLO, and Jk are qualitatively similar to those reported by us earlier [20] from the first 71 families studied. Lod scores with HLA and BF have increased; lod scores with GLO increased slightly under the recessive, 20% model (1) and decreased under the intermediate (4) and dominant (3) models. [The recessive, 50% model (2) had not been considered in the earlier paper.] The most striking drop in lod scores occurred with the Kidd blood group. Comparing interpolated lod scores along the diagonal, where θ(male) = θ(female) (table 4), the maximum Jk score dropped from 2.83 to 1.63 under model 1 and from 1.99 to 1.22 under model 3. Simultaneously, the interpolated MLE of θ rose from 0 under both models to 0.10 ± 0.12 under the recessive, 20% model, and to 0.17 ± 0.20 under the dominant, 20% model. In other words, the 29 new families reported here yield negative lod scores with Jk; considered alone, they do not continue the early trend for Jk linkage observed in our first 71 families [20].

The MZ twin concordance rate is estimated at around 20%, based on a carefully ascertained twin study [1], and we used this value as an approximation for the penetrance. We also, however, analyzed linkage under the recessive model with 50% penetrance since this model has been suggested by some researchers for an HLA-linked IDDM susceptibility locus [38]. Changing the penetrance appears to affect the estimate of θ more than the maximum lod score. The effects of penetrance on the linkage analysis will be discussed in more detail in the next section.

The lod scores reported here must be interpreted with caution. The linkage analyses assume certain genetic models for the disease, the appropriateness of which is unknown. Work done on single-locus models for IDDM has focused on inheritance at the hypothesized HLA-linked locus [5, 30, 38, 39] and therefore is not directly applicable to a possible Kidd-linked disease locus. However, by examining the four models that we did, we covered a wide range of possibilities for both possible disease loci.

Another difficulty arises in interpreting the lod scores with HLA and BF. The lod score method [25] assumes no population association between marker and trait, and this assumption is violated for both HLA and BF in our data set and in others [10–12]. However, this difficulty does not exist for GLO or Jk, since alleles at these loci are not associated with IDDM in our sample, or, for GLO, in another study [11]. Nor did any other marker exhibit strong evidence of an

association with IDDM in our data set [19]. However, we found some weak evidence for possible associations between IDDM and three markers: AMY2, GALT, and GC, as well as for a deviation from Hardy-Weinberg equilibrium at the GLO locus. (Results of individual association tests were statistically significant but did not remain so when corrected for multiple tests.) Interestingly, AMY2 (pancreatic amylase), one of the three markers just mentioned, also yielded weak positive lod scores with IDDM (table 3). AMY2 is located on the long arm of chromosome 1, 1q. Of the markers we studied, the nearest to AMY2 are Duffy (Fy), located about 14% recombination distally from AMY2 on 1q, and phosphoglucomutase-1 (PGM1), on 1p, about 26% recombination from AMY2 [40]. Close linkage with IDDM appears to be ruled out for both Duffy and PGM1 (table 3) under models 1 and 3, although loose linkage remains a possibility. Clearly, however, any hypothesis involving a second IDDM locus linked to AMY2 must remain very tentative.

The only two markers with unambiguously positive lod scores under all models were the two associated markers, BF and HLA. Since it is precisely these markers that violate the assumption of no association, these linkage results might seem questionable. On the other hand, if the positive linkage results with HLA were due only to the association, and did not represent true linkage, then we would have expected to find clear heterogeneity between families with and without the associated markers. For example, families with associated HLA haplotypes or the *DR3/4* genotype should have contributed disproportionately to the positive lod scores for HLA and BF. Analyses (B) and (C) using the Predivided-Sample Test indicate that such heterogeneity is not present (see below). This evidence supports the validity of the linkage results for HLA and BF, despite the multiple underlying assumptions required by the analysis.

Recently, a series of Minnesota families with IDDM has been analyzed for linkage with Jk, with strongly negative results [41]. These results, if combined with ours, negate our positive lod scores. For example, under the recessive model with 20% penetrance (model 1), the combined total lod scores have a maximum of approximately 0.8 at $\sim \theta = .25$.

The only reason not to combine lod scores would be if heterogeneity existed between the two studies. Heterogeneity could be due to genetic differences between the populations, errors in typing or analysis, or differences in sampling. Two factors support possible heterogeneity: (1) a heterogeneity test between the two sets of lod scores, assuming model 1, is statistically significant, with $\chi^2 = 4.7$, 1 df, $P < .05$;* and (2) the Minnesota sample exhibits an apparent association between HLA and Jk among IDDM patients [42], which we did not observe in our sample [19]. However, neither of these pieces of evidence is particularly compelling. The heterogeneity χ^2 test is based on the lod scores, and, as mentioned above, they do not have their usual statistical meaning. Moreover, the association reported in [42] was rather weak.

* This χ^2 and all other calculations are based on corrected lod scores from Minnesota (J. Barbosa, personal communication, 1983), not those in [41].

Thus, the question of linkage between IDDM and the Kidd blood group is not completely settled and hinges partly on whether our results can legitimately be combined with those in [42]. Nevertheless, the combined evidence of our diminishing lod scores and the negative Minnesota results weighs against that linkage.

The assignment of Kidd to chromosome 2 is probable, although not confirmed [24, 43]. However, since Jk is almost certainly not linked to HLA, BF, or GLO [44], its actual chromosomal assignment does not affect our conclusions.

(2) *"Two locus" Linkage Analysis/Reduced Penetrance*

The method we used to correct for the effects of a second locus on a linkage analysis represents only a rough approximation to a true two-locus linkage analysis. Under the simple model described in the APPENDIX, our penetrance reduction is accurate at the populational level, that is, for an individual chosen at random from the population. However, it allows only approximately for the effects of the second locus as it segregates within a family already known to have affected members. Thus, for "originals" in a pedigree, defined as individuals whose parents are not included in the study, our correction is accurate. However, for other family members, such as unaffected siblings of an affected child, our correction overreduces the penetrance. See the APPENDIX for more details. A fully correct two-locus linkage analysis would calculate the likelihood of the pedigree, taking both loci into account, but our correction has the virtue of simplicity. Lowering the penetrance represents a step in the right direction for two-locus linkage analysis. Moreover, since the magnitudes of the lod scores are relatively unaffected by even drastic penetrance changes (see below), our results remain valid despite the approximate nature of our two-locus correction.

Note that the "two-locus" linkage analysis, just like single-locus linkage analysis, does not address the issue of whether the genetic model being used is correct, but simply assumes the specified model.

Apart from their relevance to linkage in the presence of two-locus inheritance, the combinations of single-locus analyses also yield information on the effects of penetrance on a linkage analysis. The statistical evidence for linkage, as indicated by the magnitude of the maximum lod score, remains qualitatively the same when the specified penetrance f and the gene frequency are changed. (The gene frequency and penetrance are related; see the APPENDIX.) Table 4 reveals how lod scores vary with penetrance. In most cases, as f decreases, the maximum lod score $\hat{Z}$ drops somewhat but not markedly. For example, as f goes from .50 to .20 and then to .032 under the recessive model, $\hat{Z}$ for BF decreases from 3.03 to 2.83, and then to 2.68. In three cases—HLA under the recessive model, as f goes from .50 to .20; and GLO and Jk under the dominant model, as f goes from .20 to .032—$\hat{Z}$ increases somewhat.

The estimate of θ, $\hat{\theta}$, in contrast, is strongly affected by f. Without exception, decreasing f allows us to estimate tighter linkage (i.e., reduces $\hat{\theta}$). For example, when $f = .20$ under the recessive model, $\hat{\theta}$ with HLA is 13%; but under the "two-locus" penetrance of .032, $\hat{\theta}$ becomes 0. Clearly, unaffected relatives who share

marker types with affected family members must be counted as probable recombinants when f is high but can instead be ascribed to reduced penetrance when f is low.

The relative robustness of $\hat{Z}$ with respect to f gives us some confidence in our evidence for linkage, subject to the caveats mentioned earlier. The relative sensitivity of $\hat{\theta}$ to f, in contrast, requires that we view estimates of θ from lod score analyses with skepticism, when the genetic model and penetrance are unknown. This conclusion confirms [15, 16].

(3) *Heterogeneity*

We found no evidence for linkage heterogeneity within our data set, whether we predivided the families or used the Admixture Test.

Predivided-Sample Test. We looked for heterogeneity between simplex and multiplex families in analysis (A), table 5, because of recently reported evidence [8] that two different forms of IDDM might be appearing preferentially in the two types of families. For all four markers, $\hat{z}_1$(simplex) is 0, $\hat{\theta}_1$ is .48–.50, and $\hat{z}_2$(multiplex) equals $\hat{Z}$(overall). In other words, the simplex families do not provide negative evidence against linkage; they simply provide no linkage information at all. Thus, the results here neither support nor refute the hypothesis in [8].

In analyses (B) and (C), families were divided according to HLA characteristics, with the thought that HLA types might delineate genetically distinct subtypes of IDDM. Other studies [45, 46] have found significant differences between families with an HLA-DR3/4 proband and those without. They have speculated that these differences may indicate the action of a second locus or factor in the etiology of IDDM. However, we found no persuasive evidence for genetic heterogeneity—in the sense of two separate loci, each responsible for a different form of the disease—with HLA or BF on chromosome 6 or with Kidd on chromosome 2. The statistically significant χ^2 for GLO under analysis (C), table 5, may be attributable to the performance of multiple tests on the same data. It would be difficult to explain why there is genetic heterogeneity with GLO but not with HLA and BF, unless possession of the DR3/4 genotype suppressed recombination with GLO.

Analysis (D) represented an attempt to find a relationship between lod scores at the Jk locus and lod scores at the three chromosome 6 markers. Positive results would have indicated two forms of IDDM, one linked to Jk and the other located on chromosome 6. However, results were negative, thus supporting the idea that *if* there are indeed two loci—an HLA-linked one and a Jk-linked one—then they contribute jointly to IDDM susceptibility, in some pattern of two-locus inheritance.

Admixture Test. If two distinct subtypes of IDDM exist, and if they correlate with a breakdown of families according to some measured criterion, such as HLA-DR type, then the Predivided-Sample Test will have greater statistical power to detect the heterogeneity than does the Admixture Test. However, if the subtypes do not correlate with any available breakdown, then the Predivided-Sample approach will fail, whereas the Admixture Test may still detect the heterogeneity. In other words, if we can correctly predict what correlates with the heterogeneity,

then the Predivided-Sample Test will more likely be productive; if not, the Admixture Test is better.

The Admixture Test offers a choice of two statistics, the "H" statistic or the "lod2" score. Where evidence of linkage already existed, that is, for HLA, BF, GLO, and Jk, the H heterogeneity statistic was used. The lod2 score, in contrast, allows for the possibility that linkage may have been hidden by heterogeneity, that is, some families may be linked and others unlinked. Therefore, it was used for the remaining markers. Again, there was no evidence for heterogeneity among the lod scores, with none of the tests even approaching statistical significance.

These negative findings do not rule out all forms of genetic heterogeneity within IDDM. First, lod scores can reveal only heterogeneity between *loci*. Heterogeneity in the sense of different alleles at one locus leading to different forms of IDDM, as proposed in [2, 5], for example, would not be revealed in lod-score heterogeneity. Second, heterogeneity could exist with respect to some locus that is not linked to any of our 28 markers, or is linked to one of our less informative markers.

However, in light of the strongly negative results of almost all our heterogeneity tests, it is unlikely that heterogeneity exists corresponding to any of the four breakdowns in the Predivided-Sample Test (table 5) or involving one of the loci for which we had reasonably large sample sizes. Moreover, given the large role that an HLA-linked or -related locus apparently plays in IDDM susceptibility, and given our large number of families and our strongly statistically negative results in all heterogeneity tests involving HLA, it does seem that we can rule out a heterogeneity hypothesis where one form of IDDM is HLA-related and the other is not—unless the unrelated form represented only a small proportion (e.g., 10%) of IDDM, as suggested by [8].

CONCLUSION

In summary, we examined three issues here. First, we continue to find compelling evidence for linkage of IDDM with HLA and BF, and less strong but still positive evidence for linkage with GLO and Jk. However, if our Jk lod scores are added to the corrected scores of Dunsworth et al. [41], then the evidence for a Kidd-IDDM linkage evaporates.

If linkage between IDDM and Kidd is ruled out, the second issue, of two-locus inheritance for IDDM, still remains viable. The "second" non-HLA-linked locus need not be linked to Kidd. Two-locus IDDM inheritance has been suggested on purely theoretical grounds [21, 22], as well as on experimental ones [47]. The two-locus correction developed here did not affect lod scores markedly. The lod scores are fairly consistent when penetrance is varied, and they do not change drastically even when the genetic model is changed. The estimates of θ, in constrast, do vary markedly when the penetrance or model is modified. We reiterate the need for caution when interpreting estimates of θ from a linkage analysis with unknown genetic model.

The third issue, heterogeneity within IDDM, also continues to hold interest. We found little evidence for heterogeneity within the lod scores. Specifically, there is little evidence for two forms of IDDM based on HLA type, except for

the one result with GLO discussed above, which is possibly due to multiple tests. Moreover, no evidence exists that one form is linked to HLA and another is linked to Kidd.

ACKNOWLEDGMENTS

We express our appreciation to Rachel Rubin for her conscientious and generous assistance with numerous aspects of this study; to Maryellen C. Sparkes and Michol Crist for their patience and expertise in doing the blood typing; to Kamton Joe and Jane Chen for their expert help in computer programming; and to Drs. David Greenberg, Jerome Rotter, and M. Anne Spence for critical reading of the manuscript.

APPENDIX

DERIVATION OF "TWO-LOCUS" LINKAGE CORRECTION

(1) Define the "overall genetic penetrance" F as the probability that an individual with the appropriate genes at all relevant loci is affected. We limit ourselves to simple quasi-Mendelian models with no sporadics and no genotype-environment interaction; that is, we assume that the reduction in penetrance from Mendelian values is constant for all two-locus genotypes that are at risk. Thus, the reduction of F from unity represents purely environmental contributions to disease etiology.

(2) Consider three simple two-locus models—recessive-recessive (R-R), dominant-recessive (D-R), and dominant-dominant (D-D). Let locus 1 have two alleles, A and a, with frequencies p and q, and locus 2 have alleles B and b, with frequencies r and s. We define the R-R model such that only the genotype *aabb* can be affected, with probability F. Thus, for the R-R model, F becomes F = Prob[affected|*aabb*]. Similarly, under the D-D model, F = Prob[affected|*AABB*, *AABb*, *AaBB*, or *AaBb*]; and under the D-R model, F = Prob[affected|*AAbb* or *Aabb*], assuming locus 1 is dominant and locus 2 is recessive.

(3) Define "locus-specific penetrances" f_i, for $i = 1,2$, as the probability that an individual with an appropriate genotype at locus i is affected. For example, under the R-R model, f_1 is the probability of an *aa* individual being affected when nothing is known about his genotype at locus 2. Thus, f_1 = Prob[affected|*aa*] and f_2 = Prob[affected|*bb*] in the R-R model. Similarly, under the D-D model, f_1 = Prob[affected|*AA* or *Aa*] and f_2 = Prob[affected|*BB* or *Bb*].

(4) The locus-specific penetrances are related to each other and to the gene frequencies as follows. Let K represent the population prevalence of the disease, and let freq(i) be the frequency of the at-risk genotype(s) at locus i. Then K = freq(1) · freq(2) · F. Since f_i = Prob[affected|at-risk genotype at locus i], $f_i = K/\text{freq}(i) = F \cdot \text{freq}(j)$ where $j \neq i$. For example, for the R-R model, $k = q^2s^2F$, so $f_1 = s^2F$ and $f_2 = q^2F$. Table 1 gives K, f_1, and f_2 for each of the three models.

(5) A limiting case of a two-locus model occurs when the gene frequency at one locus, the "ubiquitous" locus, is unity, so that the two-locus system reduces to a single-locus case. In this situation, the penetrance at the ubiquitous locus equals K, the population prevalence of the disease. The ubiquitous locus contributes no linkage information about the disease, since everyone in the population, whether affected or normal, has the same allele. At the other locus, the penetrance becomes F, the overall genetic penetrance.

An example. Consider an R-R trait with overall penetrance $F = 1$. Thus, genotype *aabb* is always affected, and no other genotype is at risk. Let allele frequencies be $q = .1$ and $s = .2$. Then the population prevalence is $K = q^2s^2F = .0004$.

Assume locus 1, the A locus, is linked to a marker M, which we are analyzing for linkage. Consider either parent in a family with two normal parents, one affected,

and one unaffected child. If the linkage analysis treats the disease as single locus and fully penetrant, each parent will necessarily be assigned heterozygous status at the disease locus. However, under two-locus inheritance, the parents may well be homozygous *aa* at the *A* locus. From the "viewpoint" of locus 1, the disease has a reduced penetrance of $s^2F = .04 = f_1$; that is, only 4% of *aa* individuals are affected, although the overall penetrance F is unity. Our penetrance reduction reflects the fact that the prior probability of a parent being *aa*, as opposed to *Aa*, is relatively high.

The situation is more complex for an unaffected sibling of an affected child, and the penetrance reduction should not be so great. The probability for a random *aa* sib to be affected under the R-R model is $(25\%)F$, as opposed to s^2F, if both parents are double heterozygotes *AaBb*, and is higher for other mating types. Since our correction applies a penetrance s^2F to *all* unaffected individuals, it overcorrects in those cases.

To apply this correction to the IDDM linkage analysis, we used $F = 20\%$, based on the estimate of MZ twin concordance rate in [1], and a population prevalence $K = .005$ [30]. It is easily shown that when F and K are fixed, f_i at either locus is bounded above by F and below by K. Thus, to cover a representative range of analyses, we let f assume three values: its upper bound F, its lower bound K, and one intermediate value, $f_i = .032$ ($f_i = .032$ was chosen because it corresponds to the case where $q = s$ in the R-R model). These three f_i values translate to three distinct sets of single-locus analyses for the D-R model, but only two sets for the R-R and D-D models, due to symmetry. Table 2 shows the sets of single-locus analyses corresponding to the R-R, D-R, and D-D models.

Analyses were not performed for ubiquitous loci (those with gene frequency of unity and penetrance equal to .20, or F), since the lod scores would necessarily be uninformative and would equal zero.

REFERENCES

1. CAHILL GF: Current concepts of diabetic complications with emphasis on hereditary factors: a brief review, in *Genetic Analysis of Common Diseases: Applications to Predictive Factors in Coronary Disease*, edited by SING CF, SKOLNICK M, New York, Alan R. Liss, 1979, pp 113–129
2. ROTTER JI, RIMOIN DL: Heterogeneity in diabetes mellitus update 1978: evidence for further genetic heterogeneity within juvenile-onset insulin-dependent diabetes mellitus. *Diabetes* 27:599–608, 1978
3. IRVINE WJ, DI MARIO UP, FEEK CM, ET AL.: Insulin antibodies in relation to islet cell antibodies and HLA antigens in insulin dependent (type I) diabetes. *J Clin Lab Immunol* 1:111–114, 1978
4. CHRISTY M, GREEN A, CHRISTAU B, ET AL.: Studies of the HLA system and insulin-dependent diabetes mellitus. *Diabetes Care* 2:209–214, 1979
5. HODGE SE, ROTTER JI, LANGE KL: A three-locus model for heterogeneity in juvenile-onset insulin-dependent diabetes. *Ann Hum Genet* 43:399–412, 1980
6. CAHILL GF, MCDEVITT HO: Insulin-dependent diabetes: the initial lesion. *N Engl J Med* 304:1454–1465, 1981
7. SCHERNTHANER G: The relation between clinical, immunologic, and genetic factors in insulin-dependent diabetes mellitus, in *Genetics of Diabetes Mellitus*, edited by KOBBERLING J, TATTERSALL R, London, Academic Press, 1982, pp 99–114
8. GREENBERG DA, ANDERSON CE: The search for heterogeneity in insulin-dependent diabetes mellitus (IDDM): evidence for familial and nonfamilial forms. *Am J Med Genet* 14:487–499, 1983
9. ROTTER JI: The modes of inheritance of insulin-dependent diabetes or the genetics of IDDM, no longer a nightmare but still a headache. *Am J Hum Genet* 33:835–851, 1981

10. RAUM D, ALPER CA, STEIN R, GABBAY KH: Genetic marker for insulin-dependent diabetes mellitus. *Lancet* i:1208–1210, 1979
11. KIRK RL, THEOPHILUS J, WHITEHOUSE S, COURT J, ZIMMET P: Genetic susceptibility to diabetes mellitus: the distribution of properdin factor B (Bf) and glyoxalase (GLO) phenotypes. *Diabetes* 28:949–951, 1979
12. SVEJGAARD A, PLATZ P, RYDER LP: Insulin-dependent diabetes mellitus (joint report), in *Histocompatibility Testing 1980*, edited by TERASAKI PI, Los Angeles, UCLA Tissue Typing Laboratory, 1980, pp 638–656
13. WALSH LJ, COX DW, EHRLICH RM: The *HLA B818, BF*F1* haplotype in type 1 diabetes. *Am J Med Genet* 11:337–353, 1982
14. MCDEVITT HO, BODMER WF: HL-A, immune-response genes, and disease. *Lancet* i:1269–1275, 1974
15. HODGE SE, SPENCE MA: Some epistatic two-locus models of disease. II. The confounding of linkage and association. *Am J Hum Genet* 33:396–406, 1981
16. CLERGET-DARPOUX F: Bias of the estimated recombination fraction and lod score due to an association between a disease gene and a marker gene. *Ann Hum Genet* 46:363–372, 1982
17. LUCARELLI P, SCACCHI R, CORBO RM, ET AL.: Genetic polymorphisms in juvenile-onset diabetes. *Hum Hered* 28:89–95, 1978
18. VAGUE PL, MELIS C, MERCIER P, VIALETTES B, LASSMANN V: The increased frequency of the Lewis negative blood group in a diabetic population. *Diabetologia* 15:33–36, 1978
19. HODGE SE, ANDERSON CE, NEISWANGER K, ET AL.: Association studies between type 1 (insulin-dependent) diabetes and 27 genetic markers: lack of association between type 1 diabetes and Kidd blood group. *Diabetologia*. In press, 1983
20. HODGE SE, ANDERSON CE, NEISWANGER K, ET AL.: Close genetic linkage between diabetes mellitus and Kidd blood group. *Lancet* ii:893–895, 1981
21. THOMSON G: A two locus model for juvenile diabetes. *Ann Hum Genet* 43:383–398, 1980
22. CLERGET-DARPOUX F, BONAITI-PELLIE C, DESCHAMPS I, HORS J, FEINGOLD N: Juvenile insulin-dependent diabetes: a possible susceptibility gene in interaction with HLA. *Ann Hum Genet* 45:199–206, 1981
23. ANDERSON CE, HODGE SE, RUBIN R, ET AL.: A search for heterogeneity in insulin dependent diabetes mellitus (IDDM): HLA and autoimmune studies in simplex, multiplex and multigenerational families. *Metabolism* 32:471–477, 1983
24. SHOWS TB, MCALPINE PJ: The 1981 catalogue of assigned human genetic markers and report of the nomenclature committee. *Cytogenet Cell Genet* 32:221–245, 1982
25. MORTON NE: Sequential tests for the detection of linkage. *Am J Hum Genet* 7:277–318, 1955
26. OTT J: Estimation of the recombination fraction in human pedigrees: efficient computation of the likelihood for human linkage studies. *Am J Hum Genet* 27:588–597, 1974
27. HODGE SE, MORTON LA, TIDEMAN S, KIDD KK, SPENCE MA: Age-of-onset correction available for linkage analysis (LIPED). *Am J Hum Genet* 27:761–762, 1979
28. RAO DC, KEATS BJB, MORTON NE, YEE S, LEW R: Variability of human linkage data. *Am J Hum Genet* 30:516–529, 1978
29. OTT J: A simple scheme for the analysis of HLA linkages in pedigrees. *Ann Hum Genet* 42:255–257, 1978
30. SPIELMAN RS, BAKER L, ZMIJEWSKI CM: Gene dosage and susceptibility to insulin-dependent diabetes. *Ann Hum Genet* 44:135–150, 1980
31. GREENBERG DA: A simple method for testing two-locus models of inheritance. *Am J Hum Genet* 33:519–530, 1981
32. HLA NOMENCLATURE COMMITTEE: Nomenclature for factors of the HLA system, in *Histocompatibility Testing 1980*, edited by TERASAKI PI, Los Angeles, UCLA Tissue Typing Laboratory, 1980, pp 18–20

33. MORTON NE: The detection and estimation of linkage between the genes for elliptocytosis and the Rh blood type. *Am J Hum Genet* 8:80–96, 1956
34. KENDALL MG, STUART A: *The Advanced Theory of Statistics*, vol 2, 3rd ed. New York, Hafner, 1973, pp 234–267
35. SMITH CAB: Testing for heterogeneity of recombination fraction values in human genetics. *Ann Hum Genet* 27:175–182, 1963
36. OTT J: Counting methods (EM algorithm) in human pedigree analysis: linkage and segregation analysis. *Ann Hum Genet* 40:443–454, 1977
37. RISCH N, BARON M: X-linkage and genetic heterogeneity in bipolar-related major affective illness: reanalysis of linkage data. *Ann Hum Genet* 46:153–166, 1982
38. RUBINSTEIN P, GINSBERG-FELLNER F, FALK C: Genetics of type I diabetes mellitus: a single, recessive predisposition gene mapping between *HLA-B* and *GLO*. *Am J Hum Genet* 33:865–882, 1981
39. RAUM D, AWDEH A, ALPER CA: Bf types and the mode of inheritance of insulin-dependent diabetes mellitus (IDDM). *Immunogenetics* 12:59–74, 1981
40. COOK PJL, HAMERTON JL: Report of the Committee on the Genetic Constitution of Chromosome 1. *Cytogenet Cell Genet* 32:111–120, 1982
41. DUNSWORTH TS, RICH SS, SWANSON J, BARBOSA J: No evidence for linkage between diabetes and the Kidd marker. *Diabetes* 31:990–993, 1982
42. BARBOSA J, RICH S, DUNSWORTH T, SWANSON J: Linkage disequilibrium between insulin-dependent diabetes and the Kidd blood group Jk^b allele. *J Clin Endocrinol Metab* 55:193–195, 1982
43. LEWIS M, KAITA H, PHILIPPS S, STEINBERG AG, GIBLETT ER, ANDERSON JE: Analysis of linkage relationships of Co, Jk, and K with each other and with chromosome 2 loci ACP1 and Km. *Ann Hum Genet* 46:349–354, 1982
44. KEATS BJB, MORTON NE, RAO DC, WILLIAMS WR: *A Source Book for Linkage in Man*. Baltimore, Johns Hopkins Univ. Press, 1979, p 35
45. GREEN A, MORTON NE, ISELIUS L, ET AL.: Genetic studies of insulin-dependent diabetes mellitus: segregation and linkage analyses. *Tissue Antigens* 19:213–221, 1982
46. DUNSWORTH TS, RICH SS, MORTON NE, BARBOSA J: Heterogeneity of insulin-dependent diabetes—new evidence. *Clin Genet* 21:233–236, 1982
47. NAKAO Y, MATSUMOTO H, MIYAZAKI T, ET AL.: IgG heavy-chain (Gm) allotypes and immune response to insulin-requiring diabetes mellitus. *N Engl J Med* 304:407–409, 1981